HUMAN HEMATOPOIETIC STEM AND PROGENITOR CELL ANTIGEN AND METHODS FOR ITS USE

INTRODUCTION

Technical Field

[0001] This invention relates to antigens expressed by hematopoietic stem cells and progenitor cells and to methods of using such antigens, especially for cell separation and purification.

Background

[0002] The high turnover of mammalian blood cells requires a supply of hematopoietic stem cells that are able to give rise to other blood cell lineages. The immediate progeny of the hematopoietic stem cell are called progenitor cells, and are capable of giving rise to various cell types within one or more lineages, *i.e.* the erythroid, myeloid and lymphoid lineages. The stem cell and progenitor cell populations constitute only a small percentage of the total number of cells in bone marrow, fetal liver, *etc.* These populations are of immense interest because of their ability to repopulate the hematopoietic system.

[0003] A number of methods have been described in the literature for the purification or enrichment of hematopoietic stem cell and progenitor cell populations. There is significant commercial interest in these methods because hematopoietic progenitors have a number of clinical uses. Progenitor cell transplantation is currently used in conjunction with chemotherapy and radiation for the treatment of leukemia, breast cancer and other tumors. Frequently, autologous transplants are used to avoid the danger of graft rejection, but there is an increased risk of disease reappearance, due to the presence of tumor cells in the engrafting cell population. Transplantation of a more purified source of progenitor cells is therefore preferable.

[0004] There is also interest in the use of hematopoietic progenitor cells as a vehicle for gene therapy. Although not yet proven in the clinic, the longevity of hematopoietic stem cells and the dissemination of their progeny in the vasculature are desirable characteristics. A number of vectors, including several retrovirus and adenovirus based constructs, that can transfect hematopoietic stem cells have been described.

[0005] Proteins and other cell surface markers found on hematopoietic stem cell and progenitor cell populations are of great interest, as they are useful in preparing reagents for identification, separation and isolation of these populations and in the further characterization of these important cells. Although some antigens are now known that can be used in the identification and separation (positive and negative) of stem cells, such as (for example) the CD 34 antigen, which is found on stem cells but not on mature blood cells, there is a continued need for development of other antigens, particularly one that can simplify the identification and separation of desirable classes and subclasses of cells, especially hematopoietic stem cells and progenitor cells.

Background Literature

[0006] U.S. Patent No. 5,061,620 describes a substantially homogeneous human hematopoietic stem cell composition and the manner of obtaining such composition. Stromal cell-associated hematopoiesis is described by Paul *et al.* (1991) Blood **77**.:1723-1733. The phenotype of stem cells with rhodamine staining is discussed in Spangrude and Johnson (1990) P.N.A.S. **87**:7433-7437. Cell surface antigen expression in hematopoiesis is discussed in Strauss *et al.* (1983) Blood **61**:1222-1231 and Sieff *et al.* (1982) Blood **60**:703-713. Descriptions of pluripotential hematopoietic cells are found in McNiece *et al.* (1989) Blood **74**:609-612 and Moore *et al.* (1979) Blood Cells **5**:297-311. Characterization of a human hematopoietic progenitor cell capable of forming blast cell-containing colonies *in vitro* is found in Gordon *et al.* (1987) J. Cell. Physiol. **130**:150-156 and Brandt *et al.* (1988) J. Clin. Invest. **82**:1017-1027. The use of progenitor cells in transplantation is discussed in To *et al.* in Progenitor Threshold in Transplantation (ISBN 1-880854 17-1) pp. 15-20. Utilities for the cell compositions obtained using the methods and compositions of the invention are described in these publications, among others.

[0007] The use of high-gradient magnetic separation for the isolation of human hematopoietic progenitor cells is described in Thomas and Landsdorp (1992) in Advances in Bone Marrow Purging pp.537-544; and Kato and Radbruch (1993) Cytometry **14**:384-392. Other methods of magnetic selection for human hematopoietic progenitor cells are described in Bigas *et al.* (1992) in Advances in Bone Marrow Purging pp.545-551; Oku *et al.* (1992) in Advances in Bone Marrow Purging pp. 553-560; and Hardwick *et al.* (1992) in Advances in Bone Marrow Purging pp. 583-589. High gradient magnetic cell sorting is described in Miltenyi *et al.* (1990) Cytometry **11**:231-238. Molday, U.S. 4,452,773 describes the preparation of magnetic iron-dextran microspheres and provides a summary describing the various means of preparation of particles suitable for attachment to biological materials.

SUMMARY OF THE INVENTION

[0008] Methods and compositions are provided for the enrichment and characterization of human hematopoietic progenitor and stem cells. An antigen has been identified, referred to here as the AC133 antigen, that is present on stem cells and on progenitor cells and that can be used for the identification and/or separation of these important cells from the vast majority of cells present in a biological (or other) source of hematopoietic cells. Novel antigen compositions and reagents that react with them, such as antibodies, are provided for use in the methods of the invention and for the further investigation of hematopoietic progenitor and stem cell biology. For example, hematopoietic cells can be obtained from various sources, including fetal and adult bone marrow, cytokine mobilized peripheral blood cells, and fetal liver, and can be separated using reagents and methods of the invention.

BRIEF DESCRIPTION OF THE DRAWINGS

[0009] The invention now being generally described, the same will be better understood by reference to the following description of specific embodiments together with the figures that form part of the current specification, wherein:

[0010] Figure 1 shows a dot-plot from fluorescence activated cell sorting (FACS) analysis of fetal liver cells. The *y* axis represents cell staining with AC133 antibody

conjugated to phycoerythrin (PE). The cells were counterstained with HPCA2-FITC (anti-CD34). The numbers represent the percent of total cells that fall within the quadrants.

[0011] Figure 2 is a graph showing FACS analysis of AC133 antigen expression on phorbol myristate acetate (PMA) activated Y79.1 cells.

[0012] Figure 3 is a graph showing FACS analysis of AC133 antigen and CD34 expression on PMA activated Y79.1 cells.

[0013] Figures 4A and 4B show dot-plots from 3 color FACS analysis of the antibodies AC133, CD38 and HLA-DR on fetal liver cells. The *x* axis in Figure 4A represents HLA-DR-FITC, and the *y* axis represents cell staining with AC133-PE. The *x* axis in Figure 4B represents CD38-FITC, and the *y* axis represents cell staining with AC133-PE.

[0014] Figures 5A, 5B, 5C and 5D show dot-plots from FACS analysis of the antibodies CD38, HLA-DR, CD90 and CD117 on AC133 positive cells purified from fetal liver. In Figure 5A the *x* axis represents CD38-FITC staining, and the *y* axis represents HPCA2-PE staining. In Figures 5B, 5C and 5D, the *x* axis represents staining with HPCA2-FITC. The *y* axis in Figure 5B represents cell staining with anti-HLA-DR-PE. The *y* axis in Figure 5C represents cell staining with anti-CD90-PE. The *y* axis in Figure 5D represents cell staining with anti-CD117-PE. The numbers represent the percent of total cells that fall within the boxed gates.

[0015] Figure 6 is a gel showing the results of an immunoprecipitation with AC133 antibody and the cell lines KG1a and Y79.1. The lanes are as follows: 1) molecular weight markers; 2) a 1:5 dilution of unprecipitated KG1a lysate; 3) a 1:50 dilution of unprecipitated Y79.1 lysate; 4) KG1a lysate precipitated with AC101 antibody (CD34); 5) Y79.1 lysate precipitated with AC101 antibody; 6) KG1a lysate precipitated with HPCA2 antibody (CD34); 7) Y79.1 lysate precipitated with HPCA2 antibody; 8) KG1a lysate precipitated with 16D11 antibody (CD34); 9) Y79.1 lysate precipitated with 16D11 antibody; 10) KG1a lysate precipitated with AC133 antibody; 11) Y79.1 lysate precipitated with AC133 antibody; 12) mixed kG1a and Y79.1 lysate precipitated with a mixture of AC133 and HPCA2 antibodies; 13) KG1a lysate precipitated with 8A3 (anti-CD109) antibody; 14) KG1a lysate precipitated with 15G5 (anti-CD109) antibody.

[0016] Figures 7A and 7B are graphs showing FACS analysis of CD56 (Figure 7A) and AC133 antigen (Figure 7B) expression on PMA activated, tunicamycin treated or untreated Y79.1 cells.

[0017] Figure 8 is a dot-plot showing FACS analysis of HPCA2-PE (*y* axis) staining of AC133 magnetically purified fetal liver cells. The *x* axis represents staining with a glycophorin A-FITC conjugate.

[0018] Figures 9A and 9B are dot plots showing FACS analysis of HCPA2 staining of buffy coat peripheral blood mononuclear cells before and after AC133 magnetic separation. The *y* axis shows staining with HCPA2-PE, the *x* axis shows staining with anti-CD45 and anti-CD15-FITC conjugated antibodies.

[0019] Figure 10 is a bar graph showing the cloning efficiency of AC133 and AC101 purified cells in a clonogenicity assay.

[0020] Figure 11 is a bar graph showing the plating efficiency of AC133 purified cells and AC133 negative, CD34 positive cells.

[0021] Figure 12 is a chemical formula showing DNA and amino acid sequences for the AC133 antigen.

[0022] Figure 13 is a schematic diagram of the transmembrane and other regions of the AC133 antigen.

DESCRIPTION OF SPECIFIC EMBODIMENTS

[0023] Methods and compositions are provided that have use in the enrichment and/or characterization of human hematopoietic stem cells and/or progenitor cells. The immediate progeny of the hematopoietic stem cell, called here "progenitor" cells, are capable of giving rise to various cell types within one or more lineages. In the present invention, stem cells and/or a sub-set of progenitor cells (i.e., CFU-GM cells that are needed for short-term ingraftment) can be identified or selected through the use of reagents that specifically bind to a newly discovered antigen referred to here as the AC133 antigen (Ag) that is highly specific for these cells. The high tissue specificity of AC133 antigen expression is particularly advantageous during enrichment for highly purified progenitor cell populations. An AC133-positive cell population is highly enriched for cells that are active in assays

measuring progenitor cell activity, particularly in the CFU-GM activity. The subset of cells that is AC133 negative and CD34 positive is enriched for BFU-E activity, a measure of erythroid-committed progenitor cell activity.

[0024] Reagents that specifically bind to the AC133 antigen include without limitation physiological ligands, synthetic ligands, polyclonal antibodies, and monoclonal antibodies. An AC133 monoclonal antibody is any monoclonal antibody which interacts specifically with the AC133 cell antigen expressed on a subset of hematopoietic progenitor cells derived from human bone marrow, fetal bone marrow and liver, cord blood and adult peripheral blood. The subset of progenitor cells recognized by antibodies directed to AC133 are $CD34^{bright}$ and contains substantially all of the CFU-GM activity present in the $CD34^{+}$ subset (as well as those cells that are still stemcells collected in a collection of progenitor cells). For purposes of transplantation, cells active in CFU-GM are of particular interest because they provide for production of neutrophils. Use of an AC133 antibody allows positive immunoselection of hematopoietic progenitor cell populations, as well as the phenotypic analysis of progenitor cell populations using flow cytometry. In particular, an antibody against AC133 recognizes not just CFU-GM cells, which are needed for short-term engraftment and protection from sepsis, but also primitive long-term re-populating cells that are necessary for long-term engraftment. Cells selected for expression of AC133 antigen can then be further purified and/or separated by selection for other hematopoietic stem cell and progenitor cell markers.

[0025] As outlined below in detail, molecules of interest in the various methods of the invention include the AC133 antigen itself, reagents that specifically bind to AC133 or a fragment thereof, AC133 complexed to a ligand, an AC133-ligand complex wherein the ligand is an antibody, nucleic acid sequences encoding the AC133 antigen, and population of cells that express the AC133 antigen or any of its fragments. The AC133 antigen can be isolated from natural sources or produced using recombinant DNA technology. The nucleic acids can be cDNA, RNA, a genomic sequence, or a synthetic sequence comprising the coding sequence by itself or in conjunction with transcriptional regulatory regions and other sequences found in expression and/or cloning vectors. The AC133 Ag itself can be obtained in a purified form by isolation from cells, which can be identified as positive by AC133

antibody binding using affinity binding methods known in the art. Positive identification is available by proteolytic digestion of cell membrane proteins and comparison of sequences to the protein sequence for AC133 set out in Figure 12.

[0026] mAb AC133 is an antibody with specificity for a novel cell surface antigen that is expressed on bright $CD34^+$ cells. The antigen is expressed on a subset of hematopoietic progenitor cells derived from human bone marrow, fetal bone marrow and liver, cord blood, and adult peripheral blood. mAb AC133 can be used in a magnetic bead system to immunoselect hematopoietic progenitor cell populations, resulting in potential therapeutic benefit, as well as in the phenotypic analysis of progenitor cell populations using flow cytometric techniques. To further characterize the nature of this novel molecule, the AC133 antigen was purified by immunoaffinity chromatography. The AC133 antigen consists of a single polypeptide chain with a reduced molecular weight of about 120 kD, and comprises a glycoprotein with an about 20-kDa N-glycosidic-linked polysaccharides. The reduced AC133 antigen is recognized by mAb AC133, suggesting a linear epitope or a sugar epitope.

[0027] It will be recognized by those experienced in the field of glycoproteins that such molecules are not expected to have exactly identical sugar structures because of the enzymatic nature of sugar synthesis, which occurs without the template (i.e., messenger RNA) that exists for peptide synthesis, although similarities will certainly exist among the sugar structures in a collection of AC133 molecules because of synthesis from the same starting peptide structure. Accordingly, "AC133 antigen" refers to proteins having the peptide structure shown in Figure 12 (discussed below) with sugar structures attached at glycoslyation sites. Because of the natural variations in sugar structures, a range of molecular weights for glycosylated molecules is also to be expected and comes within the scope of the present invention. In the case of AC133 antigen, there appears to be relatively low variation in the structure and size of the attached sugar residues, compared to other known proteins. Molecular weight of AC133 is typically found to be in the range of 115 to 127 kD, regardless of the details of the experimental technique used to measure molecular weight.

[0028] The purified AC133 antigen was digested with lysyl endopeptidase to generate peptides that were isolated by reverse phase HPLC and sequenced by Edman degradation. These peptides were used to design degenerate oligonucleotides used in the polymerase chain reaction with a WERI-Rb-1 cDNA library template. This technique yielded 1.7 kB of unambiguous sequence which was then used to isolate the entire cDNA clone. This cDNA encodes a single open reading frame of 2598 nucleotides, and predicts a 865 amino acid protein with a molecular weight of 96.8 kDa, which corresponds with the about 90 kDa molecular weight found for the deglycosylated antigen. Hydrophobicity and transmembrane helicity analysis suggests the presence of five transmembrane domains, resulting in two large extracellular loops. There are a total of 8 consensus sequences for sites of N-linked glycosylation, all on the two large (260 and 290 a.a.) loops supporting our proposed structural model with two large extracellular loops and a 50 amino acid C-terminal cytoplasmic tail. A truncated version of the AC133 antigen missing the cytoplasmic tail is still recognized by mAb AC133. There are consensus sequences for a leucine zipper motif in both extracellular loops, which can be involved in receptor interaction with its physiological ligand. As shown in Figure 13, the AC133 antigen appears as a 5-transmembrane protein ("5TM protein") with an extracellular N-terminus and a cytoplasmic C-terminus.

[0029] Families of 4TM (also called tetraspan), 7TM, and 11TM proteins have been characterized in the literature. While the function of the tetraspan family is not known, the 7TM proteins are generally believed to be G-protein coupled receptors binding chemotactic agonists, and 11TM proteins represent a family of ion-channel receptors. However, a 5TM molecule has not previously been described, and the structure of the AC133 antigen differs markedly from known 7 TM family members with respect to molecular weight and size of extracellular loops. Additionally, the AC133 antigen does not share sequence homology with 4TM or 7TM proteins, while family members do share significant homology with each other, particularly within the transmembrane domains.

[0030] Short fragments of the AC133 gene are present in Genbank as EST's (expressed sequence tags), such as adult retina, pancreatic islets and fetal brain. Expression of the AC133 antigen, however, appears to be limited to primitive hematopoietic stem cells and some neural-crest-derived tissues. AC133 antigen is also expressed on NT-2

teratocarcinoma cells; however expression is lost as these cells terminally differentiate into neurons. The interaction of the physiological ligand with the AC133 antigen (receptor) can provide for intracellular signalling.

[0031] The original monoclonal antibody discovered to the AC133 antigen is one of a panel of antibodies which defines a novel antigen of approximately Mr 120,000 which is selectively expressed on $CD34^{bright}$ human hematopoietic stem and progenitor cells. $CD34^{bright}$ cells support long-term B cell lymphopoiesis and myelopoiesis in vitro and mediate T, B, myelomonocytic and megakaryocytic repopulation *in vivo*. $CD34^{dim}$ cells have failed to provide long-term hematopoietic activity *in vitro* or *in vivo*. The $CD34^{bright}$ population contains all of the primitive stem cell activity and therefore is the population of choice for further studies in hematopoietic stem cell transplantation and gene therapy. AC133 antibody provides a means for the positive selection and phenotypic analysis of hematopoietic stem cells and a subset of committed progenitor cells. The original specific antibody AC133, a murine IgG_1 antibody, was elicited from mice immunized with purified $CD34^+$ human progenitor cells. In order to determine the precise antigen phenotype of AC133 positive cells, AC133 and CD34 double positive cells were examined in fetal liver, fetal and adult bone marrow, cord blood and peripheral blood using 3 and 4 color FACS analysis. The subset recognized by AC133 antibody in all tissues are $CD34^{bright}$, $CD38^{-/+}$, $HLA\text{-}DR^{+/-}$. The $CD90^+$, $CD117^+$ and $CD109^+$ stem cell populations are included within the AC133 positive population. Typically AC133 stains 20-60% of all $CD34^+$ cells, a population which contains all the non-lineage committed $CD34^+$ population as well as CD34+ cells committed to the granulocyte/monocytic pathway. AC133 antigen expression has not been demonstrated on peripheral blood mononuclear cells, granulocytes, platelets or umbilical vein-derived endothelial cells by standard FACS procedures. FACS analysis on a panel of 50 human cell lines shows that only 2 retinoblastoma cell lines, Y79.1 and WERI-Rb-1, express AC133 antigen, along with NT-2 teratocarcinoma cells. Transplantation of AC133 positive cells into fetal sheep has demonstrated the engrafting capability of selected cells, and human cells which have homed to the fetal sheep bone marrow have been harvested and shown to engraft secondary recipients, proving the long term repopulating potential of selected cells. The AC133 gene codes for a polypeptide consisting of 865 aa with a predicted

size of 96.8 kDa. This protein has a unique structure, traversing the membrane 5 times. The AC133 antigen therefore defines a new class of mammalian 5TM membrane proteins. Together these data demonstrate that AC133 provides an alternative antigen system for the identification and separation of hematopoietic stem cells.

[0032] Antibodies that selectively bind to stem cells and/or progenitor cells are of particular interest. Antibodies to AC133 Ag can be obtained by immunizing a xenogeneic immunocompetent mammalian host (such as a murine, rodentia, lagomorpha, ovine, porcine, or bovine, host) with human hematopoietic progenitor cells. The choice of a particular host is primarily one of convenience. A suitable progenitor cell population for immunization is obtained by isolating $CD34^+$ cells from cytokine-mobilized peripheral blood, bone marrow, fetal liver, or other source of progenitor cells. The cells can be incubated with phytohemagglutinin prior to their use as an immunogen.

[0033] Immunizations are performed in accordance with conventional techniques, where the cells can be injected subcutaneously, intramuscularly, intraperitoneally, intravascularly into a host. Normally, from about 10^6 to 10^8 cells will be used, which can be divided into 1 or more injections, usually not more than about 8 injections, over a period of from about one to three weeks. The injections can occur with or without adjuvant; examples of adjuvant include complete or incomplete Freund's adjuvant, specol, and alum.

[0034] In a preferred embodiment, contralateral immunization is used, as described in the examples below. This method relies on the trafficking ability of immune lymphocytes to home to the site of antigen stimulation. The animals are pre-immunized at a localized site on one side of the body, such as a left footpad, with cells that express many immunodominant but irrelevant antigens. Various mature hematopoietic cells can be used for this purpose. The immunogen of interest is injected at a localized site on the opposite side of the animal. Lymphocytes pre-immunized with and responding to irrelevant antigens are decoyed to the left-hand draining lymph nodes, while the lymphocytes responding to the immunogen of interest will be present in the right-hand draining lymph nodes, e.g. the popliteal lymph nodes for footpad injection. This popliteal lymph node can be used as a source of cells for fusion.

[0035] After completion of the immunization schedule, the antiserum can be harvested in accordance with conventional techniques to provide polyclonal antisera specific for the surface membrane proteins of hematopoietic progenitor cells, including AC133 Ag. Lymphocytes can then be harvested from the appropriate lymphoid tissue, *e.g.* spleen or draining lymph node, and fused with an appropriate fusion partner, usually a myeloma line, to produce a hybridoma secreting a specific monoclonal antibody. Screening clones of hybridomas for the antigenic specificity of interest can be performed in accordance with conventional methods.

[0036] Of particular interest are the specific monoclonal antibody AC133 described in the Examples below; other antibodies (both monoclonal and polyclonal) that bind to the AC133 antigen, especially cross-reactive antibodies (i.e., those which bind to the same epitope, and substantially inhibit simultaneous binding); species analogs thereof; binding fragments thereof; and conjugates thereof. A deposit of a murine hybridoma cell line that expresses an antibody to the AC133 antigen was made at the American Type Culture Collection, 10801 University Boulevard, Manassas VA 20110, on April 23, 1997, and given the ATCC designation HB12346. These antibodies are capable of immunoselection for the hematopoietic subset of interest.

[0037] It is known that antibodies can be produced as a single chain instead of a normal multimeric structure. Single chain antibodies are described in Jost *et al.* (1994) J.B.C. **269**:26267-73, and in numerous other publications. DNA sequences encoding the variable region of the heavy chain and the variable region of the light chain are ligated to a spacer encoding at least about 4 amino acids of small neutral amino acids, including glycine and/or serine. The protein encoded by this fusion allows assembly of a functional variable region that retains the specificity and affinity of the original antibody.

[0038] Methods of humanizing antibodies are also known in the art. A humanized antibody can be the product of an animal having transgenic, human, immunoglobulin-constant-region genes (see for example International Patent Applications WO 90/10077 and WO 90/04036). Alternatively, the antibody of interest can be engineered by recombinant DNA techniques to substitute the CH1, CH2, CH3, hinge domains, and/or the framework residues with the corresponding human sequence (see WO 92/02190).

[0039] The use of Ig cDNA for construction of chimeric immunoglobulin genes is known in the art (Liu *et al.* (1987) P.N.A.S. **84**:3439 and (1987) J. Immunol. **159**:3521). In these techniques mRNA is isolated from a hybridoma or other cell producing the antibody and used to produce cDNA. The cDNA of interest can be amplified by the polymerase chain reaction using specific primers (U.S. Patent nos. 4,683,195 and 4,683,202). Alternatively, a library can be made and screened to isolate the sequence of interest. The DNA sequence encoding the variable region of the antibody is then fused to human constant region sequences. The sequences of human constant region genes can be found in Kabat *et al.* (1991) Sequences of Proteins of Immunological Interest, N.I.H. publication no. 91-3242. Human C region genes are readily available from known clones. The chimeric, humanized antibody can then be expressed by conventional methods.

[0040] Antibody fragments, such as Fv, $F(ab')_2$ and Fab fragments, can be prepared by cleavage of the intact antibody, *e.g.* by protease or chemical cleavage. Alternatively, a truncated gene can be designed. For example, a chimeric gene encoding a portion of the $F(ab')_2$ fragment could include DNA sequences encoding the CH1 domain and hinge region of the H chain, followed by a translational stop codon to yield a truncated antibody fragment.

[0041] Antibodies to the AC133 antigen bind to a protein that has an apparent molecular weight (under Western blot conditions from reducing SDS-PAGE gels, based on commercially available standards) of about 120 kD, and generally appears to be in the range of about 115 to 127 kD. The antibody appears to recognize a sugar epitope, as AC133 antibody cannot be immunoprecipitated from tunicamycin-treated WERI-Rb-1 cells. The AC133 antigen is expressed on a subset of $CD34^+$ cells, but is absent on endothelium and fibroblasts. Included in the population of AC133-positive cells are $HLA\text{-}DR^+$, $CD90^+$ and $CDl17^+$ progenitor cells (the antigen formerly known as $CD90^+$ is now known as $CD90^+$; both DR positive and negative as well as CD38 positive and negative cells are included in this population). This population contains substantially all of the hematopoietic stem activity present in the $CD34^+$ subset of hematopoietic cells.

[0042] Reagents that specifically bind to the AC133 antigen are not limited to antibodies. Any of numerous methods known in the art to detect the binding of one species to another can be used to assay for the presence of an AC133 antigen-binding reagent. One

universally adaptable assay involves distribution of radioactivity between soluble and solid phases can be detected using radioactively labeled test compounds and AC133 antigen attached to a solid phase. AC133 antigen can be attached, for example, to a solid phase in a column, and a tritium- or ^{14}C-labelled test compound in a physiological buffer can be passed through the column. Bound radioactivity can be detected directly on the column or by subtraction of radioactivity in the soluble phase passing through the column from the applied radioactivity. Binding affinity can be detected by measuring levels of bound radioactivity at different concentrations of test compound after allowing sufficient time for binding to equilibrate. Specificity of binding for AC133 can be detected by determining whether test compounds that bind to AC133 also bind to antigens present on mature blood cells (or other antigens of interest in a preselected assay medium). Especially preferred ligands are those that are selective for AC133 with less than 10%, preferably less than 5%, crossreactivity with any antigen present on mature blood cells. Crossreactivity can be measured by any standard technique and preferably is measured by a competitive binding assay between pure AC133 antigen, the ligand to be tested, and the suspected crossreactive antigen using a concentration of AC133 antigen and test ligand where the ligand half-saturates binding to AC133. Most preferably, crossreactivity is measured at a concentration of AC133 antigen that half saturates monoclonal antibody ATCC HB12346 when the antibody is present at a concentration of 50 ng/100 µl.

[0043] Once a reagent is identified that specifically binds to AC133, the reagent (in its radioactively labeled form, in a non-radioactive form modified to contain another label, or in certain uses in unlabeled form) can be used in various assays or biological uses that call for the binding of a reagent to AC133, such as fluorescent staining, cell separation, or cell differentiation, either *in vivo* and *in vitro*. For example, immunoselection with an antibody against AC133 provides a means of purifying hematopoietic progenitor and stem cells. The antibodies also find use in diagnostics to detect or enumerate hematopoietic progenitor cells, in dividing the CD34 positive population into functionally distinct sub-populations, in isolation of progenitor cells, and in preparation of progenitors to produce mature blood cells. Biological samples (e.g. blood or derivatives thereof, biopsies, and synovial fluid) can be assayed for the presence of cells expressing the surface molecule bound by the subject

antibodies. For example, assays can be performed on cell lysates, intact cells, or frozen sections in order to distinguish different types of cells.

[0044] The subject antibodies and other reagents that specifically bind to AC133 are useful for the preparation of substantially pure human hematopoietic progenitor and stem cells. A subset of progenitor cells can be separated from other hematopoietic cells on the basis of AC133 binding and can be further separated from each other by binding to other surface markers known in the art. Sources of hematopoietic cells include fetal or adult bone marrow; fetal liver; umbilical cord blood; and peripheral blood, particularly cytokine mobilized peripheral blood (see, for example, Campos *et al.* (1993) Leukemia 7:1409-15 and Grigg *et al.* (1993) Bone Marrow Transplant **11**, Suppl 2:23-9).

[0045] Human stem cells have been reported to have the phenotype $CD34^{bright}$; $HLA\text{-}DR^{+}$; $CD38^{dim/negative}$; $CD117(c\text{-}kit)^{dim}$; $CD90(Thy\text{-}1)^{+}$; and to lack expression of a variety of lineage specific markers, including CD3, CD4, CD7, CD8, CD14, CD15, and CD19. A negative designation indicates that the level of staining is at or below the brightness of an isotype-matched negative control. A dim designation indicates that the level of staining may be near the level of a negative stain, but may also be brighter than an isotype matched control.

[0046] Procedures for separation include magnetic separation using antibody-coated magnetic beads and affinity chromatography or "panning" using antibody attached to a solid matrix (e.g. plate). Techniques providing accurate separation include fluorescence-activated cell sorters, which can have varying degrees of sophistication, such as having multiple color channels, low angle and obtuse light scattering detecting channels, or impedance channels. Dead cells can be eliminated by selection with dyes associated with dead cells e.g., (propidium iodide, LDS). Red blood cells can be removed by (for example) elutriation, hemolysis, or Ficoll-Paque gradients. Any technique can be employed that is not unduly detrimental to the viability of the selected cells.

[0047] Conveniently, antibodies can be conjugated with labels for a number of different purposes: e.g., magnetic beads to allow for ease of separation of a particular cell type; biotin, which binds with high affinity to avidin or streptavidin; fluorochromes, which can be used with a fluorescence activated cell sorter; haptens; and the like. Multi-color

analyses can be employed with a FACS or in a combination of immunomagnetic separation and flow cytometry. Multi-color analysis is of interest for the separation of cells based on multiple surface antigens: e.g., $AC133^+$, $CD90^+$ or $CDl17^+$, $AC133^-$, or $CD34^+$. Fluorochromes which find use in a multi-color analysis include phycobiliproteins, *e.g.* phycoerythrin and allophycocyanins; fluorescein, and Texas red.

[0048] In one embodiment of the invention, an anti-AC133 antibody is directly or indirectly conjugated to a magnetic reagent, such as a superparamagnetic microparticle (microparticle). Direct conjugation to a magnetic particle is achieved by use of various chemical linking groups as known in the art. For example, antibody can be coupled to the microparticles through side chain amino or sulfhydryl groups and heterofunctional cross-linking reagents. A large number of heterofunctional compounds are available for linking to entities. A preferred linking group is 3-(2-pyridyldithio)propionic acid N-hydroxysuccinimide ester (SPDP) or 4-(N-maleimidomethyl)-cyclohexane-1-carboxylic acid N-hydroxysuccinimide ester (SMCC) with a reactive sulfhydryl group on the antibody and a reactive amino group on the magnetic particle.

[0049] Alternatively, an anti-AC133 antibody is indirectly coupled to magnetic particles. The antibody is directly conjugated to a hapten, and hapten-specific, second-stage antibodies are conjugated to the particles. Suitable haptens include digoxin, digoxigenin, FITC, dinitrophenyl, nitrophenyl, avidin, and biotin. Methods for conjugation of the hapten to a protein are known in the art, and kits for such conjugations are commercially available.

[0050] For separation or identification of stem cells or progenitor cells, an antibody is added to a hematopoietic cell sample. The amount of an anti-AC133 antibody necessary to bind a particular cell subset is empirically determined by performing a test separation and analysis. The cells and an anti-AC133 Ab are incubated for a period of time sufficient for complexes to form, usually at least about five minutes, more usually at least about 10 minutes, and usually not more than one hour, more usually not more than about 30 minutes.

[0051] The cells can additionally be incubated with antibodies or binding molecules specific for cell-surface markers known to be present or absent on hematopoietic progenitor or stem cells. For example, CD90, CD117 and HLA-DR are useful in the positive selection of stem cells. Various markers known to be absent on stem cells, such as CD3, CD4, CD8,

CD14, CD15, and CD19, can be used for negative selection. The labeled cells are separated in accordance with the specific antibody preparation. Fluorochrome-labeled antibodies are useful for FACS separation and magnetic particles for immunomagnetic selection or particularly high gradient magnetic selection (HGMS). Exemplary magnetic separation devices are described in WO/90/07380, PCT/US96/00953 and EP 438,520, herein incorporated by reference.

[0052] The purified cell population can be collected in any appropriate medium. Various media are commercially available and can be used, including Dulbecco's Modified Eagle Medium (DMEM), Hank's Basic Salt Solution (HBSS), Dulbecco's phosphate buffered saline (DPBS), RPMI, Iscove's modified Dulbecco's medium (IMDM), and phosphate buffered saline (PBS) with 5 mM EDTA, any of which can be supplemented with fetal calf serum (FCS), bovine serum albumin (BSA), or human serum albumin (HSA).

[0053] Compositions highly enriched for human hematopoietic progenitor and/or stem cells (depending on the source of cells) are achieved in this manner in a single step. The desired cells will be at or about 80% or more of the cell composition, and preferably be at or about 90% or more of the cell composition. Specific populations of interest include $AC133^+$ cells, which are characterized as $CD34^{bright}$ and $HLA\text{-}DR^{+/-}$. This population can be further selected for those cells that are $CD90^+$, $CD117^+$ and/or $CD38^{dim}$. Functionally these cells are highly enriched for CFU-GM activity and for long-term re-populating cells. Another population of interest is $CD133^-$ and $CD34^+$, which is enriched for BFU-E activity. The use of the subject antibodies for purification are advantageous over the use of CD34, because AC133 is expressed by a more restricted population of cells, thereby permitting a more enriched subset for the specific activity of interest.

[0054] Once the desired cells have been isolated, they can be propagated by growing in conditioned medium from stromal cells, co-culturing with such stromal cells, or in media comprising maintenance factors supporting the proliferation of such progenitor cells e.g., stem cell factor or combinations of interleukins. The medium employed for culturing cells is conveniently a defined enriched medium, such as IMDM or a mixture of IMDM and RPMI 1640, and will generally be composed of salts, amino acids, vitamins, 5×10^{-5} M 9-

mercaptoethanol, streptomycin/penicillin and 10% fetal calf serum, and can be changed from time to time, generally at least once to twice per week.

[0055] The subject cell compositions find use in a variety of ways. They can be used to reconstitute an irradiated host and/or a host subject to chemotherapy. By providing for maturation, proliferation and differentiation into one or more selected lineages through specific different growth factors the progenitor cells can be used as a source of committed cells. Such factors as erythropoietin, colony stimulating factors (e.g., GM-CSF, G-CSF or M-CSF), interleukins (e.g. IL-1, -2, -3, -4, -5, -6, -7, -8, -9, or -10), or the like, or stromal cells can be used to influence the growth and differentiation of progenitor cells.

[0056] The cells can also be used in the isolation and evaluation of factors associated with the differentiation and maturation of hematopoietic cells, including reagents that specifically bind to the AC133 antigen. Thus, the cells can be used in assays to determine the activity of media, such as conditioned media; to evaluate fluids for growth factor activity or involvement with dedication of lineages; or the like.

[0057] The cells can be used for the treatment of genetic diseases. Genetic diseases associated with hematopoietic cells can be treated by genetic modification of autologous or allogeneic stem cells to correct a genetic defect or treat to protect against disease, e.g., HIV. For example, diseases such as 9-thalassemia, sickle cell anemia, adenosine deaminase deficiency, recombinase deficiency, or recombinase regulatory gene deficiency can be corrected by introduction of the wild-type gene into the subject cells, either by homologous or random recombination. Alternatively, normal allogeneic progenitor cells can be transplanted. Diseases other than those associated with hematopoietic cells can also be treated, where the disease is related to the lack of a particular secreted product such as hormone, enzyme, interferon, factor, or the like.

[0058] The cells can be frozen at liquid nitrogen temperatures and stored for long periods of time, as they can be thawed and reused. The cells will usually be stored in 5% DMSO and 95% fetal calf serum. Once thawed, the cells can be expanded by use of growth factors or stromal cells associated with stem cell proliferation and differentiation.

[0059] The AC133 antigen can be obtained in substantially pure form from either natural sources or by recombinant techniques. From natural sources, the antigen-positive

cells are lysed and passed through an affinity column of anti-AC133 monoclonal antibody. Hematopoietic progenitor cells can be isolated from natural sources by conventional separation techniques, or cell lines described in the experimental section can be used as a source of antigen. The affinity-purified protein is eluted from the affinity column with an appropriate salt solution or aqueous/organic gradient, such as acetonitrile or ethanol, usually in the presence of a low acid concentration, e.g., 0.1-1 percent trifluoroacetic acid. The eluted protein is then further purified by chromatography, electrophoresis, or the like in accordance with conventional techniques.

[0060] The examples below describe the use of a monoclonal antibody to purify the AC133 antigen by affinity chromatography resulting in greater than 95% pure AC133 antigen. Peptides of such a purified preparation can be prepared and isolated for sequence analysis, as a result of which nucleic acid probes can be designed for the isolation of AC133 gene sequences. The gene sequence of AC133 set forth herein (Figure 12) allows the antigen to be obtained by recombinant techniques. For example, total RNA is isolated from cells that have been shown by antibody binding to express the targeted protein. Residual DNA is removed in accordance with conventional techniques, and the polyadenylated RNA can be purified further, for example on oligo-dT sepharose or by gel chromatography. cDNA is then prepared in accordance with conventional techniques using reverse transcriptase (see Sambrook *et al., supra* and the Examples below). The cDNA is then introduced into an appropriate cloning system, such as Σ gt11, where the cDNA is expressed. The phage plaques can then be screened using the subject antibodies, or by employing polyclonal antisera. Alternatively, a cloning system can be used which allows probing with nucleic acid sequences derived from the AC133 antigen protein sequence. The cDNA inserts are then subcloned into other vectors, as desired. The cDNA can be used for further probing of the cDNA library for a complete transcript. Alternatively, the cDNA sequence can be used to probe a genomic library to identify the genomic gene encoding the subject proteins (See, for example, Molecular Cloning: A Laboratory Manual, 2nd ed., J. Sambrook, E.F. Fritsch, T. Maniatis, CSHL, Cold Spring Harbor, NY, 1989).

[0061] DNA of the invention includes the nucleotide sequences encoding the AC133 protein or fragments thereof, as well as adjacent 5' and 3' non-coding nucleotide

sequences involved in the regulation of expression of the protein encoded by the genes, and will include up to about the length of the mature mRNA or genomic DNA. Thus, the present invention provides an isolated nucleic acid molecule, in which the molecule comprises (1) a first sequence having an amino acid coding region for AC133 as set forth in Figure 1 (SEQ ID NO:1); (2) a second sequence, wherein said second sequence is a subsequence of said first sequence and is at least 14, preferably at least 17 or 20, more preferably at least 25, nucleotides in length; (3) a third sequence in which at least one nucleotide of said first or second sequences is replaced by a different nucleotide; or (4) a fourth sequence complementary to any of said first, second or third sequences; with the proviso that (i) if said molecule is an RNA molecule, U replaces T in said sequence of said molecule, (ii) said third sequence is at least 90%, preferably at least 95%, identical to said first or second sequence, and (iii) said second sequence is not nucleotides 347-667, 1564-1696, or 2110-2386 of SEQ. ID NO:1. Also included as DNA of the invention is the corresponding genomic sequence, including introns. These non-coding sequences include terminator and polyadenylation sequences, regulatory protein binding sequences, transcriptional sequences, and the like. Molecules containing the full length AC133 cDNA sequences are useful as sources of subsequences or as starting materials for the preparation of the AC133 molecule itself.

[0062] A "subsequence" is a group of consecutive nucleotides from the cDNA sequence. Any of these sequences can be used in the identification of the presence (or absence) of the AC133 gene or of the expression of mRNA encoding the AC133 antigen. Such subsequences can be prepared by chemical synthesis from starting nucleotides (as in an automated gene synthesizer) or by biochemical manipulation of the full-length sequences (e.g., using restriction endonucleases to prepare fragments, optionally followed by (1) cleavage of terminal nucleotides and exonucleases and/or (2) size sorting and/or affinity capture to select the desired sequence). Any subsequence of the AC133 sequence described in SEQ ID No.: 1 of sufficient length to be unique among the other nucleic acids present under the conditions being used is useful as one of the two primers used in a polymerase chain reaction (PCR) amplification of all or part of the genomic AC133 gene. The length of a subsequence necessary to hybridize uniquely with the desired target sequence will vary with the particular method being used, and selection of the length is within the ordinary skill

of those who carry out routine identification of genetic material. A preferred subsequence is at least 15 nt in length, more preferably at least 18 nt, even more preferably at least 19, 20, 21, 25, or 30 nt in length up to the full length of the nucleotide sequence shown as SEQ. ID NO:1, preferably less than 200 nt in length if used as a hybridization probe or less than 50 nt in length if used as a PCR primer.

[0063] Three subsequences within the coding region of SEQ. ID NO:1 were previously recorded in Genbank as EST's of unknown function. Accordingly, these Genbank subsequences, nucleotides 347-667, 1564-1696, and 2010-2386, are not claimed as subsequences of the invention. Additionally, there are a number of EST's in Genbank from the 3' untranslated region of SEQ. ID NO:1, also of unknown function, specifically in the regions covered by nucleotides 2684-3332 and 3408-3804. Subsequences from these two regions are not claimed as part of the present invention. Longer subsequences of the entire sequence shown as SEQ. ID NO:1 that contain one or more of the Genbank sequences, as well as subsequences of any length that include part of one or more Genbank sequence but also contain newly identified nucleotides set forth in SEQ. ID NO:1, are considered to be part of the present invention.

[0064] The nucleic acid compositions of the subject invention can be genomic or cDNA sequences encoding all or a part of the subject protein. Fragments can be obtained of the cDNA or genomic sequence by chemically synthesizing oligonucleotides in accordance with conventional methods, such as by restriction enzyme digestion or by PCR amplification. For the most part, fragments will be of at least 12 nt, more usually at least 18 nt, or one of the other lengths described above. Preferred fragments will include a functional epitope. The sequence providing for a functional epitope can be determined by expression of the sequence, and assaying for reactivity of the expression product with specific antibodies by conventional immunoassay.

[0065] Exemplary amino acid and DNA sequences of the invention are set forth in SEQ ID No.: 1 and 2 below. Standard abbreviations for nucleotides and amino acids are used in this specification. Polypeptides derived from the natural AC133 antigen are particularly preferred embodiments of the invention, although variations based on the specific sequences of these polypeptides are also parts of the present invention. In its broader

aspects the invention (as it pertains to polypeptides per se) includes any polypeptide selected from the group consisting of (1) a first amino acid sequence of AC133 as set forth in SEQ ID NO: 2; (2) a second amino acid sequence wherein the second sequence is a subsequence of the first sequences and is at least 6, preferably 8, more preferably 10, amino acids in length; or (3) a third sequence in which at least one amino acid of the first or second sequences is replaced by a different amino acid, with the proviso that the amino acid replacement is a replacement of one acidic residue for another, one basic residue for another, one non-polar residue for another, one uncharged polar residue for another, or one aromatic residue for another, with the proviso that the third sequence is at least 90%, preferably 95%, identical to the first or second sequence.

[0066] Two amino acid sequences are homologous if there is a partial or complete identity between their sequences. For example, 85% homology means that 85% of the amino acids are identical when the two sequences are aligned for maximum matching. Gaps (in either of the two sequences being matched) are allowed in maximizing matching. Gaps of 5 or less are preferred with 2 or less being more preferred. Alternatively and preferably, two protein sequences (or polypeptide sequences derived from them of at least 30 amino acids in length) are homologous as this term is used herein, if they have an alignment score of more than 5 (in standard deviation units) using the program align with the mutation data matrix and a gap penalty of 6 or greater (Dayhoff, M. O., in Atlas of Protein Sequence and Structure, 1972, vol. 5, National Biomedical Research Foundation, pp. 101-110, and supplement 2 to this volume, pp. 1-10). The two sequences or parts thereof are more preferably homologous if their amino acids are greater than or equal to 50% identical when optimally aligned using the align program.

[0067] Minor amino acid variations from the natural amino acid sequence sets forth in SEQ ID No.: 2 are contemplated; in particular, conservative amino acid replacements are contemplated. Conservative replacements of those that take place within a family of amino acids that are related in their side chains. Genetically encoded amino acids are generally divided into four families: (1) acidic, aspartate, glutamate; (2) basic: lysine, arginine, histidine; (3) non-polar: alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan; and (4) uncharged polar: glycine, asparagine, glutamine, cysteine,

serine, threonine, tyrosine. Phenylalanine, tryptophan and tyrosine are sometimes classified jointly as aromatic amino acids. For example, it is reasonable to expect that an isolated replacement of a leucine with an isoleucine or valine, an aspartate with a glutamate, a threonine with a serine, or a similar replacement of an amino acid with a structurally related amino acid will not have a major effect on the binding properties of the resulting molecule, especially if the replacement does not involve an amino acid as a binding site involved in the interaction of AC133 or its derivatives with a reagent that binds specifically to AC133. Whether an amino acid change results in a functional peptide can readily be determined by assaying the specific binding properties of the polypeptide derivative.

[0068] As shown in Figure 13, there are a number of regions having different functions in the peptide structure of AC133. These regions can be described (beginning with the amino terminus) as an extracellular N-terminus, a first transmembrane region, a first cytoplasmic loop, a second transmembrane region, a first extracellular loop, a third transmembrane region, a second cytoplasmic loop, a fourth transmembrane region, a second extracellular loop, a fifth transmembrane, and a cytoplasmic C-terminus. Approximate sizes of the regions are shown in Figure 13, with best estimates of the amino acids present in the different regions being as follows: extracellular N-terminus, aa 20-107; first transmembrane region, aa 107-126; first cytoplasmic loop, aa 127-157; second transmembrane region, aa 158-179; first extracellular loop, aa 180-435; third transmembrane region, aa 436-454; second cytoplasmic loop, aa 455-480; fourth transmembrane region, aa 481-503; second extracellular loop, aa 504-792; fifth transmembrane, aa 793-816; and cytoplasmic C-terminus, aa 817-865. There appears to be a cleavable signal sequence (aa 1-19) at the amino terminus of the encoded peptide; this sequence is not included as part of the regions shown in Figure 13 but will be present in synthetically produced AC133 peptides.

[0069] Also shown in Figure 13 are the approximate locations of short peptide segments (P1-P4) that were identified to verify the structure of the AC133 antigen and of glycosylation sites (indicated by a "Y" at the point of attachment). Figure 12 also shows the glycosylation sites (which are boxed in the amino acid sequence) and transmembrane regions (which are underlined). Two glycosylation sites overlap (NNTS, which consists of an

overlapping NNT and NTS) and are shown by a larger box with dashed lines indicating the individual consensus glycosylation sites.

[0070] The DNA sequences can be obtained in substantial purity and can be obtained as an isolated molecule other than a sequence of an intact chromosome. Usually, the DNA will be obtained substantially free of other nucleic acid compounds, generally being at least about 50%, usually at least about 90% pure and are typically "recombinant", i.e., flanked by one or more nucleotides with which they are not normally associated with on a natural chromosome.

[0071] The DNA sequences are used in a variety of ways. They can be used as probes for identifying related surface proteins in the same or other species. The DNA can also be used to identify cells or organs that are expressing the subject genes. Techniques in which one probes cells for the presence of particular nucleotide sequences, particularly as DNA, mRNA or cDNA, are well-established in the literature and do not require elaboration here. Conveniently, mRNA can be isolated free of DNA, and by using reverse transcriptase and PCR with specific primers, the subject cDNAs of interest of can be expanded, separated on gel electrophoresis and then probed using Southern blotting or sequencing. Other techniques can also find use.

[0072] Homologous sequences are those with substantial sequence similarity to AC133 antigen sequences included within the subject invention, i.e., at least 80%, preferably at least 90%, more preferably at least 95%, sequence identity with the nucleotide sequence of the subject DNA sequence. Sequence similarity is calculated based on a reference sequence, which can be a subset of a larger sequence, such as a conserved motif, coding region, or flanking region. A reference sequence will usually be at least about 18 nt long, more usually at least about 30 nt long, and can extend to the complete sequence that is being compared. Such homologous nucleic acid sequences will be detected by hybridization under low stringency conditions, for example, at 50° C and 10XSSC (0.9 M NaCl/0.09 M sodium citrate) and remain bound when subject to washing at 55°C with 1XSSC.

[0073] For expression, the DNA sequences can be inserted into an appropriate expression vector, where the native transcription initiation region can be employed or an exogenous transcriptional initiation region. The promoter can be introduced by recombinant

methods *in vitro*, or as the result of homologous integration of the sequence into a chromosome. A wide variety of transcriptional initiation regions are known for a wide variety of expression hosts, where the expression hosts can involve prokaryotes or eukaryotes, particularly *E. coli*, *B. subtilis*, mammalian cells, such as CHO cells, COS cells, monkey kidney cells, lymphoid cells, particularly human cell lines, and the like. Generally a selectable marker operative in the expression host will be present. The promoter can be operably linked to the coding sequence of the genes of interest so as to produce a translatable mRNA transcript. Expression vectors have convenient restriction sites located near the promoter sequence so as to provide for the insertion of nucleic acid sequences encoding heterologous proteins. The promoters in suitable expression vectors can be either constitutive or inducible. Expression vectors for the production of fusion proteins, where the exogenous fusion peptide provides additional functionality, such as increased protein synthesis, stability, reactivity with defined antisera, or an enzyme marker, e.g., β-galactosidase, are of particular interest.

[0074] Expression cassettes can be prepared comprising the transcription initiation region, which can be constitutive or inducible, with or without an enhancer sequence, including the endogenous or heterologous enhancer sequence, the AC133 gene or fragment thereof, and a transcriptional termination region, optionally having a signal for attachment of a poly A sequence. The gene can be genomic, including the native introns, or cDNA gene, or portion thereof. Of particular interest is the use of sequences which allow for the expression of functional epitopes, usually at least about 24 nucleotides in length, more usually at least about 48 nucleotides in length, and up to the complete open reading frame of the gene.

[0075] After introduction of the DNA, the cells containing the construct can be selected by means of a selectable marker, the cells expanded and then used for expression. Where secretion is desired, a signal peptide can be joined to the sequence encoding the subject proteins or fragments thereof, whereby the protein will be expressed, translocated through the cell membrane, and processed to remove the signal peptide.

[0076] The expression cassettes can be introduced into a variety of vectors, where the vectors will normally be characterized by the ability to provide selection of cells comprising the expression vectors. The vectors can provide for extrachromosomal

maintenance, particularly as plasmids in bacteria or viruses in eukaryotic cells, or for integration, particularly in mammalian cells. Where extrachromosomal maintenance is desired, an origin sequence will be provided for the replication of the plasmid, which can be a low- or high-copy plasmid. A wide variety of markers are available for selection, particularly those which protect against toxins, more particularly against antibiotics. The particular marker which is chosen will be selected in accordance with the nature of the host, where in some cases, complementation can be employed with auxotrophic hosts, e.g., yeast. Introduction of the DNA construct can be by any convenient means, e.g., calcium-precipitated DNA, electroporation, fusion, transfection, or infection with viral vectors.

[0077] The following examples are offered by way of illustration and not by way of limitation.

EXAMPLES

[0078] *Generation of the AC133 Monoclonal Antibody by Contralateral Immunization.* Five New Zealand Black (NZB) mice were inoculated a total of seven times over a twenty day period, via the footpad route, with purified CD34 positive human progenitor cells, which had been pre-incubated with phytohemagglutinin (PHA) (Gibco/BRL).

[0079] Mice were pre-immunized on Day -3 in the left hand footpad with cells that express many immunodominant but irrelevant antigens. In this case peripheral blood mononuclear cells (PBMC) were used as an irrelevant cell, as they express many antigens such as Class I HLA antigens, HLA-DR, CD15, CD26, CD29, CD31, CD36, CD44, CD45, CD58, etc., which are also expressed on hematopoietic stem cells. On day 0 PBMC are reinjected into the left footpad, and purified stem cells are injected into the right hand footpad. PBMC and purified stem cells are pre-incubated with PHA for ten minutes and washed with PBS prior to injection. Progenitor cells were isolated from a leukaphoresis pack of a cytokine mobilized donor using immunomagnetic beads. This treatment provides non-specific adjuvant effects, and obviates the need adjuvants such as Freund's. Mice are given a total of 5-8 such injections at three days intervals.

[0080] On day 21, one day after the last injection, the mouse right hand popliteal lymph nodes were removed. A lymphocyte suspension was prepared, and the cells fused to

SP2/0 Ag14 myeloma cells using a modification of the method originally described by Kohler and Milstein (1975) Nature **256**:495-497. Cells were plated on 96 well plates in DMEM + 20% fetal calf serum, with 10^{-4} M hypoxanthine and 2 μg/ml azaserine (Buck *et al.* (1984) in Monoclonal Antibodies and Functional Cell Lines Kennet *et al.* eds., Plenum Press, New York pp.275-309). On day 10, visible hybridoma colonies were apparent. Supernatants (s/n) from hybridoma containing wells were screened for binding to a fetal liver cell preparation containing up to 15% CD34+ cells, using a 2 color flow cytometry assay. Binding of mouse Ig containing s/n to the test cells was traced with rat anti-mouse Ig-conjugated to phycoerythrin (IgPE) and counterstained with a known mouse anti-CD34 antibody (AC101) conjugate. Figure 1 shows the results from this two color FACS analysis using AC133 supernatant. AC133 is shown to stain only the bright CD34 positive cells in the fetal liver preparation. AC133 hybridoma cells were shown to secrete an IgG1/kappa antibody. The cells were expanded in culture and stocks frozen in liquid nitrogen. AC133 cells were subcloned by limiting dilution analysis and a series of positively secreting subclones were also frozen in liquid nitrogen.

[0081] *Antibody purification and conjugation.* AC133 cells were initially grown as an ascites tumor in nude mice, with collection of antibody-rich ascites fluid. More recently AC133 cells have been grown to very high density in a hollow fiber culture device (Cellmax QUAD artificial capillary system, Cellco Inc., Germantown, MD). Pure IgG antibody was prepared from hollow fiber cultures or from ascites fluid by Protein A chromatography. Pure antibody was stored in 0.01M phosphate buffered saline (PBS) with 0.01% sodium azide at 4°C. This pure antibody stock was used to prepare fluorescein isothiocyanate (FITC) (Wofsy *et al.* (1980)in Selected Methods in Cellular Immunology, Mishell and Shiigi eds., W.H. Freeman and Co., San Francisco. pp.294-295), phycoerythrin (PE) (Hardy (1986) in Handbook of Experimental Immunology, Weir *et al.*, eds. Blackwell Scientific Press, Oxford. p.31), or magnetic bead conjugates, according to standard protocols.

[0082] *AC133 expression on normal tissues and cell lines.* Using standard FACS staining procedures, there was no detectable staining of peripheral blood mononuclear cells, granulocytes or platelets, or human umbilical vein endothelial cells with AC133 antibody. Examination of a panel of human cell lines by FACS analysis (data shown in Table 1)

showed that only three cell lines tested, the retinoblastoma cell lines Y79.1 and WERI-Rb-1 and the teratocarcinoma cell line NT-2, expresses detectable levels of AC133 antigen.

Table 1. AC133 Expression on Human Cell Lines

Cell Line	Cell Type	AC133
8402	T cell line (CD34+)	-
8866	B-LCL	-
AZ676	breast carcinosarcoma	-
BJAB	N. American Burkitts' lymphoma	-
BT474	breast tumor	-
BT549	breast tumor	-
BT20	breast tumor	-
CaCL74-36	melanoma	-
Daudi	B-LCL-	-
Du4475	breast tumor (CD34+)	-
HEL92.1.7	erythroleukemia	-
HL-60	promyelocytic leukemia	-
HPB-ALL	acute lymphocytic leukemia	-
HS-R	myeloma (EBV+)	-
HT1080	fibrosarcoma	-
HT29	colon adenocarcinoma	-
IM-9	B-LCL	-
JM	T cell line	-
Jurkat	T cell line	-
KG1a	acute myelogenous leukemia (CD34+)	-
KG1	acute myelogenous leukemia (CD34+)	-
KG1a.5	acute myelogenous leukemia (CD34+)	-
K562	erythroleukemia	-
MOLT-4	T cell line	-
MCF-7	breast tumor	-
Raji	B-LCL	-
RPMI 8226	myeloma	-
SK HEP-1	hepatoma	-
U937	histiocytic lymphoma	-
WERI-Rb-1	retinoblastoma	+
Y79.1	retinoblastoma	+
NT-2	teratocarcinoma	+

[0083] Activation of Y79.1 cells with PMA was found to increase the expression of AC133 antigen (shown in Figure 2). However, PMA activation of several other cell lines, or

PHA activation of human PBMC was unable to induce the expression of AC133 antigen (data are shown in Table 2). AC133 antigen expression was not detectable on any of the $CD34^+$ cell lines tested. This finding, along with the lack of CD34 expression on the Y79.1 cell line (shown in Figure 3), excludes the possibility that AC133 is directed to the CD34 antigen. AC133 antigen expression is limited to primitive stem and progenitor cells, unlike the CD34 antigen, which is also expressed on endothelium and fibroblasts (Krause *et al.* (1996) Blood **87**:1-13).

[0084] The AC133 antigen is expressed on the $CD34^{bright}$ population of human progenitor cells isolated from fetal and adult bone marrow, fetal liver, cord blood, leukaphoresis (LP) packs and LP packs from cytokine mobilized donors. Typically it stains 30-50% of all $CD34^+$ cells in these populations.

Table 2. Activation of Cell Lines

Addition of PMA at 1 ng/ml for:						
	0 hr	24 hr	48 hr	72 hr	96 hr	144 hr
Y79.1	dim+	+	+	+	+	+
KG1a	-	-	-	-	-	N/A
K562	-	-	-	-	-	N/A
HEL 92.1.7	-	-	-	-	N/A	N/A
Jurkat	-	-	-	N/A	N/A	N/A
8402	-	N/A	N/A	-	N/A	N/A
Addition of PWM at 10 μg/ml for:						
	0 hr	24 hr	48 hr			
Y79.1	dim +	+	+			
KG1a	-	-	-			

[0085] *Phenotyping of AC133 positive cells.* Phenotyping of AC133 and CD34 double positive cells was accomplished using 2 and 3 color FACS analysis, employing a panel of conjugated antibodies directed to cell surface structures known to be expressed on

progenitor cells. Fetal liver, fetal and adult bone marrow, cord blood and peripheral blood were all used to determine the precise phenotype of AC133 positive cells. AC133 cells found in all of these tissues are $CD34^{bright}$, $CD38^{-/+}$ and HLA-$DR^{-/+}$. The data are shown in Figure 4. The CD90 $(Thy1)^{+}$ and CD117 $(c\text{-}kit)^{+}$ stem cell populations are included within the AC133 positive population, as shown in Figure 5. In a series of experiments performed with AC133 immunomagnetically purified fetal liver cells, CD38-FITC conjugated antibody stained 74.5% of the AC133 purified cells, while 24.8% were CD38 negative. As expected, HLA-DR stained the majority of the cells (81.14%). CD90 is shown to stain 27.4% of the test cells, while CDl17 stained 90%. It is generally believed that primitive (repopulating) hematopoietic stem cells have the phenotype of $CD34^{bright}$, $CD38^{dim/neg}$, HLA-DR^{+}, $CD117^{dim}$ and $CD90^{+}$. Thus, the AC133 antibody recognizes a phenotypically important population of human hematopoietic progenitor cells.

[0086] *Immunoprecipitation of the AC133 antigen.* Immunoprecipiation experiments showed that the AC133 antigen has a molecular weight of 120 kD. Biotin (Pierce) labeled, activated Y79.1 and Weri-RB-1 cells were solubilized with lysis buffer: 2.5% Brij (Sigma), 25 mM Tris-HCl, pH8.0, 125 mM NaCl, 2.5 mM EDTA, 2.2 Tg/ml Aprotinin (Sigma) and 1 mM PMSF (Sigma). The lysates were incubated with AC133 and control antibodies after preclearing. Immunocomplexes were collected on *Staphylococcus aureus* cells (CalBiochem) and heated for five minutes at 95°C in SDS-PAGE sample buffer with 1% 2-mercaptoethanol. Immunoprecipitates were resolved by SDS-PAGE and transferred to nitrocellulose membranes (Novex). Visualization was accomplished using streptavidin linked to horseradish peroxidase (HRP) (Amersham) and the Supersignal CL-HRP substrate system (Pierce). CD49d, CD71 and CD98 were used as controls, and their expected bands of 133 kD, 92 kD and 80/40 kD were observed in the corresponding lanes. The immunoprecipitation with AC133 showed a distinct band corresponding to a molecular weight of 120 kD. This band was absent in the samples that were immunoprecipitated with the anti-CD34 antibodies AC101, HPCA1 and HPCA2, indicating that CD34 is not expressed in the Y79.1 cell line. This is consistent with the FACS data.

[0087] The data from a further experiment is shown in Figure 6, where biotin labeled Y79.1, as well as KG1a cells, confirmed the AC133 molecular weight data by

comparing CD34 and AC133 precipitates on the same gel. In this experiment, CD34 and Y79.1 antigens were precipitated from biotinylated KG1a ($CD34^+$) and Y79.1 lysates in adjacent lanes. The results clearly demonstrate that 1) each antibody precipitates its own distinct antigen, and 2) that the molecular weight of these two antigens is distinctly different, being 110 and 127 kD, respectively. In control lanes 6 and 8, HPCA2 and 16D11 (anti-CD34) precipitate a band of 110 kD from KG1a lysate, but do not precipitate anything from Y79.1 lysate (lanes 7 and 9). AC133 precipitates a 120 kD protein from the Y79.1 lysate (lane 10), but nothing from the KG1a lysate in lane 11. In lane 12, KG1a and Y79.1 lysates were mixed, and AC133 Ag and CD34 were co-precipitated. The results show that the two antigens are of different molecular weights.

[0088] *AG133 magnetic bead conjugation.* Purified AC133 antibody was conjugated to magnetic amino-dextran beads using a standard protocol for 4-(N-maleimidomethyl)-cyclohexane-1-carboxylic acid N-hydroxysuccinimide ester (SMCC). AC133 antibody was added to SMCC activated beads at 5 Tg per OD_{450} unit, and incubated at room temperature for two hours. The reaction was stopped by the addition of 9-mercaptoethanol and NEM. The conjugate was purified over two columns in the presence of a magnetic field, and eluted. The concentration was adjusted to OD_{450}=10, and OPG was added for stabilization. The conjugate in PBS and 0.1% sodium azide was filtered through a 0.2 Tm filter, and stored at 4°C.

[0089] *Separation of human hematopoietic progenitor cells with AC133 magnetic bead conjugate.* AC133 direct magnetic bead conjugates were prepared and tested on buffy coat PBMCs, fetal liver WCl, fetal bone marrow and adult bone marrow. Figure 8 shows the FACS dot plot of fetal liver cells purified with AC133 bead conjugate using the miniMACS system and stained with glycophorin A-FITC and HCPA2-PE. The starting material contained 7.4% $CD34^+$ cells, following AC133 purification, greater than 90% of AC133 purified cells were bright CD34+. Figure 9 shows that AC133 magnetic conjugate was also very effective in enriching $CD34^+$ cells from a buffy coat which contained about 0.26% $CD34^+$ cells. The final purified population was 64% positive for CD34, as shown by HPCA2-PE staining. This ability to separate cells in a magnetic purification system enables further study of the functional and phenotypic properties of AC133.

[0090] *Clonogenic potential of AC133 positive cells.* AC133 magnetic beads selected cells purified from leukaphoresis packs were tested in clonogenicity assays using a commercially available kit (Stem Cell Technologies, Vancouver, B.C.). By providing a controlled growth environment utilizing recombinant human growth factors this culture assay identifies the major colony forming units (CFU) within a CD34 positive cell population. It provides information on the composition of progenitor cell populations, with respect to the relative percentages of cells committed to a particular lineage specific differentiation. Typically in peripheral blood derived $CD34^+$ cell populations BFU-e (burst forming units-erythroid), and CFU-GM (colony forming units-granulocyte macrophage) are the predominant colonies recognized, being present at a 3:1 ratio. Figure 10 shows the results from a typical clonogenicity experiment comparing AC133 and CD34 purified cells obtained from a split leukaphresis pack. Colonies obtained with unfractionated control cells are typically predominantly BFU-E (29.34%), with a smaller number of CFU-GM (5.14%). CD34 purified cells show a similar distribution with 23.3% BFU-E, and 5.58% CFU-GM. In contrast, AC133 purified cells show a different pattern, with 13.1% BFU-E and 10.2% CFU-GM. Calculations show that 58% of CFU-GM were recovered in the AC133 purified fraction, while only 13% of BFUEs were recovered.

[0091] Figure 11 shows the results from a similar clonogenicity assay obtained following AC133 immunomagnetic purification. In this experiment, AC133 cells were positively selected, and then CD34 positive cells were positively selected from the AC133 negative flow-through. This design allowed the direct comparison of $AC133^+$ cells with $CD34^+$ but $AC133^-$ cells from the same donor. The results indicate that 93.8% of the CFU-GM progenitors were recovered in the AC133 positive fraction, the remaining 6.2% being recovered from the CD34+/$AC133^-$ fraction. Conversely, the $CD34^+AC133^-$ fraction contained 78.0% of the BFU-e progenitors, while the remaining 22.0% were contained in the $AC133^+$ fraction.

[0092] The above experimental results rule out the possibility that an anti-AC133 antibody is an antibody to Fc receptors, or that an anti-AC133 antibody binds to stem cells via Fc receptor uptake. Further experiments rule out the possibility that AC133 antibody staining is due to free PE. AC133 antibody does not behave like an antibody to RTK, a

receptor tyrosine kinase, TIE, a tyrosine kinase that contains immunoglobulin-like domains and growth factor homology domains and which is expressed in vascular endothelial cells and hematopoietic cells. AC133 antibody also does not behave like an antibody to P-glycoprotein, a 170 kD multi-drug resistance product which is also expressed in hematopoietic cells.

[0093] We have shown that AC133 antibody recognizes an antigen expressed only on bright CD34+ cells in bone marrow, fetal liver and peripheral blood. This antibody and its antigen do not match the molecular weight or distribution of any known CD antigen. Apart from stem cells, AC133 antibody has been shown to react with a human retinoblastoma cell line that is negative for CD34 expression. AC133 antigen is, in addition, not expressed on a number of CD34+ cell lines.

[0094] It is evident from the above results that the subject invention provides for a novel antigen found on primitive stem cells and a subset of hematopoietic progenitor cells, as well as antibodies that specifically bind to the antigen. Expression of the antigen is highly tissue specific. It is only detected on a subset of hematopoietic progenitor cells, and is present on substantially all cells that are active in the CFU-GM assay. This highly specific distribution of AC133 antigen makes it exceptionally useful as a reagent for isolating and characterizing human hematopoietic progenitor and stem cells.

[0095] *Purification and characterization of the AC133 antigen.* The purification and characterization of the AC133 antigen, as well as the isolation of a cDNA clone is described here. Protein and nucleic acid sequence analysis of this molecule indicate that the AC133 antigen is the first described member of a new class of transmembrane receptors, having 5 transmembrane domains with little if any homology to known G-protein coupled 7 transmembrane family members.

[0096] Antibody AC133 was prepared and purified as described above and conjugated to CNBr activated sepharose. CNBr activated sepharose was purchased from Pharmacia (Alameda, CA), and mAb AC133 affinity resin was prepared per the manufacturer's procedure using a 25 minute ligand coupling reaction. The COS-7 and the WERI-Rb-1 retinoblastoma cell lines were obtained from American Type Culture Collection

(Rockville, MD). Custom primers were synthesized by Operon Technologies (Alameda, CA).

[0097] *Purification of the AC133 Antigen.* The AC133 antigen was isolated from 96 hour PMA activated Y79 retinoblastoma cells (commercially available, for example, from ATCC). Cells ($2x10^9$) were washed with PBS and lysed in 0.125M NaCl, 25mM Tris pH 8, 0.005% NaN_3, 2.5mM EDTA, and 2.5% Brij 99/96 (2:1) detergent containing 1.0mM phenylmethyl sulfonylfluoride (PMSF) and a 1/1000 dilution of a 2.2 mg/ml solution of aprotinin containing 4.1 trypsin inhibitor units per mg (Sigma). Cells were vortexed intermittently for 5 minutes at room temperature and then left on ice for 20 minutes. Cell nuclei and debris were removed by centrifugation at 12,000X G for 10 minutes. Lysate supernatant was filtered through a 0.2 μM filter prior to loading onto 0.5mL mAb AC133 affinity column equilibrated in wash buffer (0.125 M NaCl, 25 mM Tris pH 8.0, 0.01% NaN_3, 2.5 mM EDTA, 0.1% Brij). The column was washed extensively with wash buffer and the antigen was eluted in 50 mM ethanolamine pH 11.5, 0.1% Brij, 0.01% NaN_3. The pH was immediately adjusted to neutral with HCl. Passage of the antigen eluate over a 300 μl bed volume DEAE column equilibrated in wash buffer removed many of contaminating proteins, and a second affinity chromatography step using an AC133 antibody column as described above resulted in >95% pure AC133 antigen amenable to proteolysis and protein sequence analysis. The purity and identity of AC133 antigen was confirmed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and Western analysis (Towbin, H., T. Staehelin, and J. Gordon (1979) PNAS 76,4350-4354; Towbin, H. and J. Gordon (1984) J. Immunol. Meth 72:313-340).

[0098] *Endoglyconase treatment of the purified AC133 antigen.* One microgram of AC133 antigen was resuspended in 50μl water and 125μl 0.1 M 2-mercaptoethanol and 0.5% SDS. The protein was denatured at 100°C for 5 minutes. Denatured mixture (35 μl) was added to each of 5 tubes, together with 25μl 0.5 M Tris pH 8, 10 μl water, 10μl 10% NP-40. 0-0.1 unit PNGase F (Sigma) was added to each tube, and the tubes were incubated at 30°C overnight. Deglycosylated antigen was visualized on a silver stained SDS-polyacrylamide gel.

[0099] *Lysyl endopeptidase digestion of the AC133 Antigen and isolation of peptides.* AC 133 antigen was precipitated from 1.4 mL of 2 μg/mL affinity column eluate by the addition of TCA to 10%. The precipitated dry protein was suspended in 25 μL of solution digest buffer (8M urea, 400 mM Tris pH 7.8), to which 5 μl of 45 mM DTT was added and the mixture incubated at 50°C for 15 min. After cooling to room temperature, 5 μl of 100 mM iodoacetamide was added and the mix was incubated for an additional 15 minutes. Distilled water (70 μl) was added, diluting the urea to 2 M, and 2 pmol of the lysyl endopeptidase, LysC (commercially available from Wako Chemicals, USA), was added. The digestion was carried out at 37°C for 24 hours. Peptides were isolated by HPLC separation on a VYDAC narrowbore C18 reverse phase column with a 4-32% acetonitrile gradient in 0.1% trifluoroacetic acid (TFA).

[00100] *Protein sequence analysis of AC133 antigen peptides.* N-terminal sequence analysis was determined using Edman chemistry (Edman, P., Begg, G. (1967) Eur. J. Biochem. 1, 80-91; Huwick, R.M., Hunkapillar, M.W., Hood, L.E., and Dreyer, W.J. (1987) J. Biol Chem. 256, p. 7990) on an Applied Biosystems 477A or 473A liquid pulse protein sequenator. PTH-Amino acids were separated on a Brownlee C-18 reverse phase column (2.1mm x 22 cm) at 55°C in buffer A (3.5% tetrahydrofuran with addition of 2 to 4% ABI Premix Buffer concentrate from Applied Biosystems to buffer B (acetonitrile), with a 12-36% buffer B linear gradient over 18 min, followed by a 13 min. isocratic period at 36% B.

[00101] *Isolation and protein sequencing of the AC133 antigen.* The 120kD AC133 antigen was isolated by immunoaffinity chromatography from a retinoblastoma cell line, Y79, which was PMA activated for 96 hours prior to harvest. Sequential affinity chromatography and DEAE chromatography were utilized to generate >95% pure AC133 antigen by SDS-PAGE and silver staining and the identity of the purified molecule as the AC133 antigen was confirmed by Western blotting. De-glycosylation of the antigen with PGNase F to remove N-linked sugar shows that approximately 30 kD of the molecular weight is due to glycosylation. Repeated initial attempts to sequence the N-terminus of the AC133 antigen failed, suggesting that this protein is amino-terminally blocked. However, digestion of the purified antigen with lysyl endopeptidase followed by reverse phase HPLC,

yielded four peptide sequences with lengths of 12-16 amino acids. Searches of the major protein and nucleic acid databases with the peptide and resulting degenerate oligonucleotide sequences indicated that the AC133 antigen could not be identified with any described molecules. (The amino acid sequence has now been deduced from cDNA cloning and is shown in Figure 12.)

[00102] *cDNA Cloning.* Total RNA was isolated from WERI-Rb-1 retinoblastoma cells (available from the American Type Culture Collection; Rockville, MD) and poly A^+ RNA was prepared using the Poly A^+ Tract System (Promega Corp., Madison, WI). cDNA was synthesized (Guebler, U. and B.J. Hoffman (1983) Gene 25:263) using superscript reverse transcriptase (GIBCO BRL, Gaithersburg, MD) and an oligo dT primer. The blunted cDNA was ligated to nonself-complimentary Bst XI adaptors and gel purified to remove unligated adaptors and small fragments. The linkered cDNA was then ligated into the pcDNA -I expression vector (Invitrogen, San Diego, CA) and electroporated into *Escherichia coli* strain MC1061/P3 (Dower, W.J. (1990) Genetic Engineering V. 12 Edited by J.K. Seflow, Plenum Press, New York 275-295. (Electroporation of Bacteria: a general approach to genetic transformation); Ausubel, F.M., R. Brent, R.E. Kingston, D.D. Moore, J.G. Seidman, J.A. Smith and K. Struhl. 1987-1994 Current Protocols in Molecular Biology. John Wilest Sons; N.Y.). WERI-Rb-1 library cDNA (100ng/reaction) was used as a PCR template with 100 pmol each degenerate sense and antisense primers designed from the protein sequence of four AC133 antigen peptides. PCR reactions were carried out in buffer (50 mM KCl, 10 mM Tris pH 9, 0.1% Triton X-100, 1.5 mM MgCl, 0.2 mM (each) dNTP's) with 5 units of Taq DNA polymerase per reaction (Promega Corp, Madison, WI). Amplification was carried out in an MJ research (data) instrument as follows: 92°C for 1 min, 55-37°C for 1 min, 72°C for 3 min, 35 cycles. After amplification the reaction mixtures were run on 1% agarose gels, and unique bands not appearing in the individual primer controls were gel purified and cloned into pCR 2.1 using a TA Cloning Kit (Invitrogen, San Diego, CA). The 5' and 3' ends of the gene were isolated by hemi-specific PCR with nested sets of AC133 antigen gene specific primers and library specific primers. Twenty cycles of single-stranded PCR were performed with each gene specific primer in a 50 µl reaction volume with 100 ng of the library cDNA and 10 pmol each primer in PCR reaction buffer (

described above) with 5 units of Taq polymerase. An aliquot (10 μl) of this reaction mix was removed and used as template for a second, 35 cycle, PCR reaction using both the gene specific primer and the library specific primer. An aliquot (5 μl) of this PCR reaction mix was then used for another 35 cycles of reaction using nested library and gene specific primers. Bands corresponding to the 5' and 3' ends of the gene were gel purified and cloned into pCR 2.1. Overlapping cDNA clones were sequenced by the dideoxy chain reaction using fluorescent dye terminators and an ABI sequencer (Applied Biosystems, Foster City, CA.)

[00103] *Isolation of a cDNA clone of the AC133 antigen.* To isolate the cDNA for this protein, a cDNA library was prepared from the WERI-Rb-1 retinoblastoma cell line that expresses approximately 10-fold more AC133 antigen than PMA activated Y79 cells. Degenerate primers were used in low stringency PCR reactions with the library to yield a 1.7 kb fragment that contained the correct sequence of peptide 3 at the 5' end and the correct sequence of peptide 4 at the 3 prime end. Additionally, the sequence of peptide 2 was found within the fragment in the correct reading frame. Hemi-specific PCR with gene specific primers and library specific primers yielded additional 1.2kB and 2kB fragments corresponding to the 5' and 3' ends of the gene and overlapping with the initial 1.7 kB clone.

[00104] Sequencing of the three partial clones yielded a 4 kB cDNA containing an open reading frame of 3.0 kB, but also containing a 128 bp intron that appears to be associated with the poly A^+ version of the gene, and does not contain eukaryotic consensus splice sequences. To isolate an intact stem cell derived clone without the intron, $AC133^+$ stem cells were isolated from fetal liver utilizing a magnetic conjugate of mAb AC133 and the Miltenyi magnetic separation system (Miltenyi Biotech, GMBH). Total RNA was isolated from these cells, and used as a template for RT-PCR reactions. Primers designed to span the intron generate a single 582 bp fragment with the poly A+ derived cDNA template, but generate a single 454 bp fragment without the intron from total RNA in $AC133^+$ cell lines (Fig 3), suggesting that the spliced mRNA is the major product within the total RNA pool. RT-PCR was utilized to generate cDNA clones originating before the start methionine and containing the complete cDNA sequence. The full length cDNA encoding AC133 antigen predicts a protein of 863 amino acids with a molecular weight of 96.8kD (Fig. 4).

Hydrophobicity analysis of the sequence (Fig. 5) and transmembrane helix algorithms indicate that the protein spans the cell membrane a total of five times (Fig. 6) predicting the presence of two large (255 and 280 amino acids) extracellular loops and a C-terminal cytoplasmic tail. Other structural features suggested by the protein sequence include leucine zipper motifs in both of the putative large extracellular loops and six consensus sequences for N-glycosylation.

[00105] *Expression of the AC133 antigen in transfected COS-7 cells.* AC133 positive cells ($1x10^7$) were isolated from fetal liver as described above. Total RNA was isolated using RNAzol (Gibco BRL, Gaithersburg, MD) as described (Chomczynski, P. and Sacchi, N. (1987) Anal. Biochem. 162, 156). RT-PCR was performed using the Promega Access RT-PCR system (Promega Corp, Madison, WI) with 10 ng total RNA template and primers directed before the start methionine and after the stop codon. The 2.8 kb band corresponding to the coding region of the gene was cloned into the Invitrogen directional eukaryotic TA cloning vector (pCR 3.1) containing the CMV promoter. Subconfluent COS-7 cells (available from the ATCC, Rockville, MD) were transfected with 5 μg of cloned DNA by electroporation and incubated for 48 hours prior to FACS analysis. Transfected COS-7 cells were stained with 50 ng/100 μl test mAb AC133-PE, and analyzed with a Becton Dickenson (San Jose, CA) FACS scan.

[00106] *Expression of the AC133 antigen in COS cells.* COS cells transfected with the AC133 antigen gene were stained with mAb AC133-PE and analyzed by FACS (Fig 7). Cos cells transfected with the AC133 antigen gene stain brightly with mAB AC133-PE, however, untransfected cells, cells transfected with empty vector or the gene for CD-8 do not stain with this antibody.

[00107] *AC133 expression in various lymphoid and non-lymphoid cell lines.* The presence of AC133 antigen transcript in a variety of cell lines was assessed by Northern analysis. Northern blot analysis was performed by using Clontech (Palo Alto, CA) multiple tissue northern blots, and by resolving RNA samples on a 1% agarose-2M formaldehyde gel and capillary blotting overnight into nylon membrane. Total RNA was isolated with Tri Reagent, and 15μg was loaded per lane. Staining of the blot with methylene blue was used to

monitor RNA concentrations. An 800bp EcoRI fragment of the cDNA was labelled with 32P-dCTP by random priming and used as a probe.

[00108] The presence of AC133 antigen transcript in a variety of cell lines was assessed by Northern analysis. A 4.4kB mRNA transcript was detectable in WERI-Rb-1 cells as well as Y79 cells and MACS-isolated AC133+ fetal liver cells. While expression of the AC133 antigen is enhanced in Y79 cells upon PMA activation, the corresponding mRNA appears to be downregulated. In normal hemtopoetic tissue, the AC133 antigen message is detectable in fetal liver, and weakly detectable in adult bone marrow as expected due to the fact that AC133+ cells in these tissues are in a minority. The AC133 antigen transcript was also noted in non-lymphoid tissues, particularly in pancreas, kidney, and placenta. Weaker signals were observed for liver, lung, brain, and heart. This is in contrast to immunohistochemical staining of paraffin tissue sections were AC133 antigen expression was detectable only in bone marrow.

[00109] In a similar manner, other antibodies have been developed that are specific for the AC133 antigen. The following table shows antibodies, immunogens, isotypes, and cross blocking for a panel of such antibodies.

Antibody	Immunogen	Isotype	AC133 cross blocking
AC133	HSC	IgG1 kappa	+++
AC139	WERI-Rb-1	IgG1 kappa	+++
AC140	WERI-Rb-1	IgG1 kappa	+/-
AC141	WERI-Rb-1	IgG1 kappa	-
AC142	WERI-Rb-1	IgG1 kappa	ND

[00110] All publications and patent applications cited in this specification are herein incorporated by reference as if each individual publication or patent application were specifically and individually indicated to be incorporated by reference.

[00111] Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, it will be readily apparent to those of ordinary skill in the art in light of the teachings of this invention that certain changes and modifications can be made thereto without departing from the spirit or scope of the appended claims.